1. ESTADÍSTICA DESCRIPTIVA

En general las observaciones se denotan por $X_1,......,X_n$. Una vez clasificadas se tendrá las modalidades $x_1,..........,x_k$ siendo $k \leq n$. Se utilizará la notación:

n_i: frecuencia absoluta de la modalidad x_i
f_i: frecuencia relativa de la modalidad x_i
N_i: frecuencia absoluta acumulada de la modalidad x_i
F_i: frecuencia relativa acumulada de la modalidad x_i
$(l_{i-1},l_i]$: intervalo i-ésimo para datos agrupados
x_i: usualmente es la modalidad i-ésima, pero para datos agrupados en intervalos, denota la marca de clase o punto medio, es decir, $x_i=(l_{i-1},l_i]/2$
a_i: para datos agrupados en intervalos, denota la amplitud del intervalo i-ésimo, es decir,
$a_i=l_i-l_{i-1}$

1.- Medidas de posición

a) Media

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\sum_{i=1}^{n} n_i x_i = \sum_{i=1}^{k} f_i x_i$$

$\overline{X} = A + \sum_i f_i d_i$ donde $d_i=x_i-A$

$\overline{X} = A + c\sum_i f_i u_i$ donde $d_i=c.u_i$

b) Media ponderada

$$\overline{X} = \frac{w_1 X_1 + w_2 X_2 + ...+ w_k X_k}{w_1 + w_2 + ...+ w_k} = \frac{\sum wX}{\sum w}$$

c) Media geométrica

$$\overline{X}_G = \prod_{i=1}^{n}(X_i)^{1/n} = \prod_{i=1}^{n}(x_i^{n_i})^{1/n}$$

d) Media armónica

$$\overline{X}_H = \frac{n}{\sum_{i=1}^{n} 1/X_i} = \frac{n}{\sum_{i=1}^{k} n_i / x_i} = \frac{1}{\sum_{i=1}^{k} f_i / x_i}$$

e) Media cuadrática

$$\overline{X}_C = \sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{k} n_i x_i^2} = \sqrt{\sum_{i=1}^{k} f_i x_i^2}$$

Se verifica: $\overline{X}_H \leq \overline{X}_G \leq \overline{X} \leq \overline{X}_C$

f) Mediana

$Q_2 = M_e = l_{i-1} + \dfrac{\frac{n}{2} - N_{i-1}}{n_i} . a_i$ con Q_2 cuartil segundo

l_{i-1}: es el extremo inferior de la clase mediana
a_i: es la amplitud de la clase mediana
N_{i-1}: es la frecuencia acumulada de la clase anterior a la clase mediana
n_i: es el número de observaciones en la clase mediana

g) Cuartiles. Datos agrupados en intervalos

$Q_1 = l_{i-1} + \dfrac{\frac{n}{4} - N_{i-1}}{n_i} . a_i$ con Q_1 cuartil primero

l_{i-1}: es el extremo inferior de la clase en la que se superan la cuarta parte (25%) de las observaciones
a_i: es la amplitud de la clase mencionada
N_{i-1}: es la frecuencia acumulada de la clase anterior a la clase citada
n_i: es el número de observaciones en la clase Q_1.

$Q_3 = l_{i-1} + \dfrac{\frac{3n}{4} - N_{i-1}}{n_i} . a_i$ con Q_3 cuartil tercero

l_{i-1}: es el extremo inferior de la clase en la que se superan las tres cuartas partes (75%) de las observaciones
a_i: es la amplitud de la clase mencionada
N_{i-1}: es la frecuencia acumulada de la clase anterior a la clase citada
n_i: es el número de observaciones en la clase Q_3.

h) Moda

$M_o = l_{i-1} + a_i \cdot \frac{h_{i+1}}{h_{i-1} + h_{i+1}}$ con $h_i = n_i/a_i$ y $h_{i-1} = n_{i-1}/a_{i-1}$

$M_o = l_{i-1} + c \cdot \frac{n_{i+1}}{n_{i-1} + n_{i+1}}$ para amplitud constante de intervalo $h_{i-1} = n_{i-1}/c$ y $h_{i+1} = n_{i+1}/c$

2.- Medidas de dispersión

a) Recorrido

$R = x_{max} - x_{min}$

b) Rango intercuartílico

$IQR = Q_3 - Q_1$

c) Desviación media absoluta a la media

$$D.M. = \sum_{i=1}^{k} f_i \cdot | x_i - \overline{X} |$$

d) Varianza

$$V_x = \sigma^2 = \sum_{i=1}^{k} f_i \cdot (x_i - \overline{X})^2 = \sum_{i=1}^{k} f_i \cdot x_i^2 - \overline{X}^2$$

e) Cuasi-varianza

$$S_x^2 = \frac{\sum_{i=1}^{k} n_i (x_i - \overline{X})^2}{k - 1}$$

f) Desviación típica

$$\sigma_x = \sqrt{V_x}$$

g) Fórmulas abreviadas de la varianza

$V_x = V_d = \sum_i f_i \, . \, d_i^2 - (\sum_i f_i . d_i)^2$ con $x_i = d_i + A$

$V_x = c^2 . [\sum_i f_i . u_i^2 - (\sum_i u_i . f_i)^2]$ con $x_i = A + c . u_i$

h) Coeficiente de variación

$$c.v. = \frac{\sigma}{\overline{x}}$$

i) **Medidas de dispersión en torno a la mediana**

$$D = \sum_{i=1}^{k} f_i . | x_i - M_e |$$

MEDA=mediana $|x_i - M_e|$

j) Tipificación de variables

$$Z = \frac{x - \overline{X}}{\sigma_x}$$

3.- Medidas de forma

a) Coeficiente de asimetría de Pearson

$A_s = \dfrac{\overline{X} - M_o}{\sigma}$

$A_s = 0$, distribución simétrica
$A_s < 0$, distribución asimétrica o sesgada a la izquierda
$A_s > 0$, distribución asimétrica o sesgada a la derecha

b) Coeficiente de asimetría de Fisher

$g_1 = \dfrac{\mu_3}{(\sigma_x)^3}$ donde μ_3 es el momento centrado de orden 3, es decir,

$\mu_3 = \sum_{i=1}^{k} f_i.(x_i - \overline{X})^3$

$g_1 = 0$, distribución simétrica
$g_1 < 0$, distribución asimétrica o sesgada a la izquierda
$g_1 > 0$, distribución asimétrica o sesgada a la derecha

c) Coeficiente de apuntamiento de Fisher

$g_2 = \dfrac{\mu_4}{\sigma^4} - 3$

$g_2 = 0$, distribución mesocúrtica
$g_2 < 0$, distribución platicúrtica
$g_2 > 0$, distribución leptocúrtica

4.- Momentos

a) En torno al origen

$m_r = \frac{\sum_{i=1}^{n} X_i^r}{n}$. Observemos que $m_1 = \overline{x}$ y $m_2 = \overline{x}^2$.

b) En torno a la media o centrales

$\mu_r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^r}{n}$ Observemos que:

$\mu_1=0$
$\mu_2=\sigma^2$
$g_1=\mu_3/\sigma^3$
$g_2=(\mu_4/\sigma^4)-3$

$\mu_2=m_2-m_1^2$
$\mu_3=m_3-3m_2m_1+2m_1^3$
$\mu_4=m_4-4m_3m_1+6m_2m_1^2-3m_1^3$

5.- Datos atípicos-outlier y atípicos-extremos

$LI=Q_1-1,5(IQR)$ $FI=LI-1,5(IQR)$
$LS=Q_3+1,5(IQR)$ $FS=LS+1,5(IQR)$

Considerar como valores atípicos ***outlier*** los situados entre LI y FI o entre LS y FS y como atípicos ***extremos*** aquellos fuera del intervalo (FI,FS)

6.- Regresión y correlación

Línea recta: $Y=a_0+a_1X$
Parábola: $Y=a_0+a_1X+a_2X^2$
Curva de grado n: $Y=a_0+a_1X+a_2X^2+ ....+a_nX^n$
Curva exponencial: $Y=ab^X$
Curva geométrica: $Y=aX^b$

Recta de mínimos cuadrados: $y - \overline{y} = \frac{S_{xy}}{S_x^2}(x - \overline{x})$ donde $\frac{S_{xy}}{S_x^2}$ =Covarianza/ Varianza de x

Error típico: $S_e = \sqrt{\frac{\sum(y_i - y_i^*)^2}{N}}$ donde y_i* es el valor de y para los valores estimados con la recta de mínimos cuadrados

$y_i - \overline{y} = (y_i - y_i^*) + (y_i^* - \overline{y})$, variación total=variación no explicada+variación explicada

$\sum (y_i - \overline{y})^2 = \sum (y_i - y_i^*)^2 + \sum (y_i^* - \overline{y})^2$, es decir, SCT(S_y^2)=SCE(S_e^2)+SCR(S_R^2)

Coeficiente de correlación: $r = \pm\sqrt{\frac{S_R^2}{S_y^2}}$ varia entre –1 y +1, también $S_e = S_y\sqrt{1 - r^2}$. En el caso lineal toma la expresión:

$r = \frac{s_{xy}}{s_x s_y}$ (coefiente de correlación producto-momento o de Pearson) donde s_{xy} representa la covarianza de las variables x e y y s_x s_y son las desviaciones típicas de las variables.

$S_x^2 = \frac{\sum_i x_i^2 n_{i.}}{N} - \overline{x}^2$ donde $n_{i.}$ representa la frecuencia absoluta marginal de la fila i

$S_y^2 = \frac{\sum_j y_j^2 n_{.j}}{N} - \overline{y}^2$ donde $n_{.j}$ representa la frecuencia absoluta marginal de la columna j

$S_{xy} = \frac{\sum_i \sum_j x_i y_j n_{ij}}{N} - \overline{xy}$ donde n_{ij} es la frecuencia absoluta de la fila i, columna j

2. DISTRIBUCIONES DE PROBABILIDAD

1.- Concepto clásico de probabilidad:

P(A)= $\lim_{N \to \infty} \frac{n}{N}$

P(A)=$\frac{Casos\ favorables}{Casos\ posibles} = \frac{a}{N}$

2.- Propiedades de la probabilidad:

P(∅)=0

0≤P(A)≤1

$A \subset B \Rightarrow P(A) \le P(B)$

P(A∪B)=P(A)+P(B)-P(A∩B)

P(A^c)=1-P(A) donde A^c es el suceso complementario de A. En algunas ocasiones el complementario de A se representa por A’

Sucesos incompatibles o mutuamente excluyentes:
$A_i \cap A_j = \emptyset$ para todo i≠j

3.- Probabilidad condicionada:

$P(B / A) = \frac{P(A \cap B)}{P(A)}$ siempre que P(A)>0 o también P(A∩B)=P(A)P(B/A)

Si los sucesos son independientes: P(A/B)= P(A) o también P(A∩B)=P(A)P(B)

4.- Teorema de Bayes:

$$P(S_i / A) = \frac{P(S_i)P(A/S_i)}{\sum_{j=1}^{n} P(S_j)P(A/S_j)}$$

$P(S_i)$ se la denomina **probabilidad a priori**
$P(S_i/A)$ se le denomina **probabilidad a posteriori** (probabilidad sabiendo lo que ha ocurrido posteriormente)
$P(A/S_j)$ se les denomina **verosimilitudes**

5.- Probabilidad total:

$P(A)=\sum_{i=1}^{n} P(A/S_i)\, P(S_i)$ con S_i sucesos mutuamente excluyentes, partición de S y $A \cap S_i \neq \emptyset$

6.- Funciones de probabilidad o de densidad y de distribución:

a) Variable aleatoria discreta:

Función de cuantia o de probabilidad f(x):

$f(x_i)=P(X=x_i)$

Condiciones: $f(x_i) \geq 0$; $\sum_i f(x_i) = 1$

Función de distribución F(x):

$F(x)=P(X \leq x)$

Condiciones: $F(-\infty)=0$; $F(\infty)=1$

$$F(x)=\sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq x} f(x_i) = F(x_{i-1}) + f(x_i) \;\; ; \;\; f(x_i) = F(x_i) - F(x_{i-1})$$

$$P(x_i < X \leq x_j) = \sum_{k=i+1}^{j} f(x_k) = F(x_j) - F(x_i)$$

b) Variable aleatoria continua

Función de densidad f(x)

Condiciones: $f(x_i) \geq 0$; $\int_{-\infty}^{\infty} f(x)dx = 1$

$P(x_1<X<x_2)=\int_{x_1}^{x_2} f(x)dx$

Función de distribución F(x)

$F(x)=P(X<x)$

$P(X\leq x)= F(x)=\int_{-\infty}^{x} f(t)dt$

$P(x_1<X\leq x_2)= F(x_2)-F(x_1)=\int_{x_1}^{x_2} f(x)dx$

$F(-\infty)=\int_{-\infty}^{-\infty} f(x)dx = 0$;$F(\infty)=\int_{-\infty}^{\infty} f(x)dx = 1$

$F'(x)= f(x)$ donde F'(x) es la derivada de la función de distribución y f(x) es la función de densidad.

7.- Esperanza matemática:

- en variables discretas: $E[X]=\sum_{i=1}^{n} p_i x_i$
- en variables continuas: $E[X]=\int_{-\infty}^{\infty} xf(x)dx$

8.- Varianza y desviación estandar:

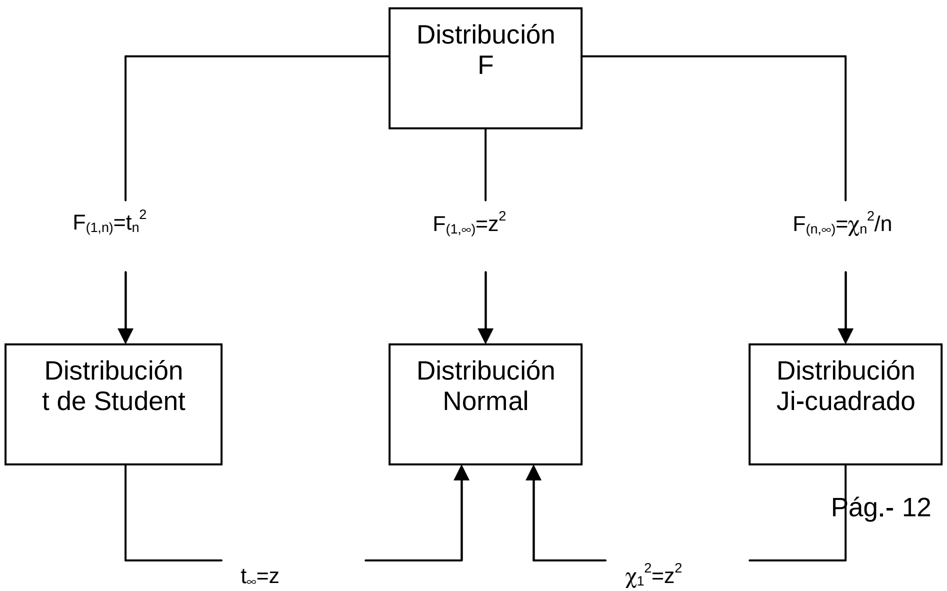
- varianza: $V(X)=\sigma^2=E[X-\mu]^2= E(X^2)-\mu^2$
    - en variables discretas: $V(X)=\sigma^2=E[X-\mu]^2=\sum_i (x_i-\mu)^2 p_i$
    - en variables continuas: $V(X)=\sigma^2=E[X-\mu]^2=\int_{-\infty}^{\infty} (x_i-\mu)^2 f(x)dx$
- desviación estandar: $\sigma=\sqrt{V(X)}$

9.- Distribuciones de probabilidad:

Distribución de probabilidad	Función de probabilidad	Parámetros	Media o esperanza matemática	Varianza
Binomial	$P(X = k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!(n-k)!} p^k q^{n-k}$	n,p $0\leq p\leq 1$	np	npq
Poisson	$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$	$\lambda>0$	λ	λ
Normal	$g(y)=\frac{1}{\sigma\sqrt{2\pi}} e^{-1/2(y-\mu)^2/\sigma^2}$	$-\infty<\mu<\infty$ $\sigma>0$	μ	σ^2
χ^2 de Pearson	$f(x;n)=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$ $\chi^2(n)$ suma de cuadrados de n, N(0,1)	n	n	2n

T de Student	f(x)= $\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}[1+\frac{x^2}{n}]^{-\frac{n+1}{2}}$ t(n)= $\frac{\lambda_0}{\sqrt{\frac{\chi^2(n)}{n}}}$ con λ_0--->N(0,1)	n	0	$\frac{n}{n-2}$
F de Snedecor	f(x)= $\frac{m^{\frac{m}{2}} n^{\frac{n}{2}}\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} x^{\frac{m}{2}-1}(n+mx)^{-\frac{m+n}{2}}$ F(x)= $\frac{n}{m}\frac{\chi^2(m)}{\chi^2(n)}$	n,m	$\frac{n}{n-2}$ si n>2	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ si n>4

Relación entre la distribución normal, ji-cuadrado, t de Student y F de Snedecor.



3. INTERVALOS DE CONFIANZA ASOCIADOS A DISTINTOS PARÁMETROS

Parámetro a estimar	Condiciones	Estimador	Distribución	Intervalo
Media	N(μ,σ) σ^2 conocida	$\bar{x} = \frac{\sum_i x_i}{n}$	Normal	$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ población infinita $\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ población finita
Media	N(μ,σ) σ^2 desconocida n>30	$\bar{x} = \frac{\sum_i x_i}{n}$	Normal	$\bar{x} \pm z_{\alpha/2} \frac{\hat{s}}{\sqrt{n}}$ población infinita $\bar{x} \pm z_{\alpha/2} \frac{\hat{s}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ población finita
Media	N(μ,σ) σ^2 desconocida n<30	$\bar{x} = \frac{\sum_i x_i}{n}$	t_{n-1}	$\bar{x} \pm t_{n-1,\, \alpha/2} \left(\hat{s} / \sqrt{n} \right)$
Proporción	n>30 0,1<p<0,9 ó np>5	$\hat{p}$ proporción muestral n. $\hat{p}$ ->B(n,p) $\frac{\hat{p}-p}{\sqrt{\frac{\hat{p}\hat{q}}{n}}} \approx N(0,1)$	Normal	$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}}$ población infinita $\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n}} \sqrt{\frac{N-n}{N-1}}$ población finita
Varianza	N(μ,σ) σ^2 conocida n<100	$\frac{(n-1)\hat{s}^2}{\sigma^2}$	Ji-cuadrado	I= $\left(\frac{(n-1)\hat{s}^2}{\chi^2_{\alpha/2;(n-1)}} ; \frac{(n-1)\hat{s}^2}{\chi^2_{1-\alpha/2;(n-1)}} \right)$

Parámetro a estimar	Condiciones	Estimador	Distribución	Intervalo
Varianza	$N(\mu,\sigma)$ σ^2 conocida $n>100$	$\frac{(n-1)\hat{s}^2}{\sigma^2}$	Ji-cuadrado	$I=\left(\frac{(n-1)\hat{s}^2}{\chi^2_{\alpha/2;(n-1)}};\frac{(n-1)\hat{s}^2}{\chi^2_{1-\alpha/2;(n-1)}}\right)$ $\chi^2_{(1-\alpha/2);(n-1)}=\frac{(\sqrt{2(n-1)-1}-z_{\alpha/2})^2}{2}$ $\chi^2_{(\alpha/2);(n-1)}=\frac{(\sqrt{2(n-1)-1}+z_{\alpha/2})^2}{2}$
Diferencias de medias	$N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$ independientes $n_1 \geq 30$ y $n_2 \geq 30$ σ^2_1 y σ^2_2 conocidas	$(\overline{x}_1-\overline{x}_2)$	Normal	$(\overline{x}_1-\overline{x}_2)\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$
Diferencias de medias	$N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$ independientes $n_1 \geq 30$ y $n_2 \geq 30$ σ^2_1 y σ^2_2 desconocidas	$(\overline{x}_1-\overline{x}_2)$	Normal	$(\overline{x}_1-\overline{x}_2)\pm z_{\alpha/2}\sqrt{\frac{\hat{s}_1^2}{n_1}+\frac{\hat{s}_2^2}{n_2}}$
Diferencias de medias	$N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$ independientes $n_1\leq 30, n_2\leq 30$ σ^2_1 y σ^2_2 desconocidas pero iguales	$(\overline{x}_1-\overline{x}_2)$	$t_{n_1+n_2-2}$	$(\overline{x}_1-\overline{x}_2)\pm t_{\alpha/2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\sqrt{\frac{(n_1-1)\hat{s}_1^2+(n_2-1)\hat{s}_2^2}{n_1+n_2-2}}$
Diferencias de medias	$N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$ independientes $n_1\leq 30, n_2\leq 30$ σ^2_1 y σ^2_2 desconocidas no iguales	$(\overline{x}_1-\overline{x}_2)$	t_m	$(\overline{x}_1-\overline{x}_2)\pm t_{\alpha/2}\sqrt{\frac{\hat{s}_1^2}{n_1}+\frac{\hat{s}_2^2}{n_2}}$

				$m=\dfrac{\left(\dfrac{\hat{s}_1^2}{n_1}+\dfrac{\hat{s}_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{\hat{s}_1^2}{n_1}\right)^2}{n_1+1}+\dfrac{\left(\dfrac{\hat{s}_2^2}{n_2}\right)^2}{n_2+1}}-2$
Diferencias de medias datos relacionados	Normales σ_d=conocido	d=x-y	Normal	$\bar{d}_0 \pm z_{\alpha/2}\dfrac{\sigma_d}{\sqrt{n}}$
Diferencias de medias datos relacionados	Normales σ_d=desconocido, n>30	d=x-y	Normal	$\bar{d}_0 \pm z_{\alpha/2}\dfrac{\hat{s}_{d_0}}{\sqrt{n}}$
Diferencias de medias datos relacionados	Normales σ_d=desconocido, n<30	d=x-y	t_{n-1}	$\bar{d}_0 \pm t_{\alpha/2,(n-1)}\dfrac{\hat{s}_{d_0}}{\sqrt{n}}$
Diferencias de proporciones	Normales, muestras grandes ($n_1p_1,n_1q_1,n_2p_2,n_2q_2>5$)		Normal	$(\hat{p}_1-\hat{p}_2)\pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$
Razón de varianzas de dos Poblaciones normales	$N(\mu_1,\sigma_1)$ y $N(\mu_2,\sigma_2)$ independientes $n_1 \geq 30$ y $n_2 \geq 30$ σ^2_1 y σ^2_2 conocidas	$\dfrac{\hat{s}_1^2}{\hat{s}_2^2}\dfrac{\sigma_2^2}{\sigma_1^2}=F_{(n_1-1).(n_2-1)}$	F de Fisher	$\left(\dfrac{\dfrac{\hat{s}_1^2}{\hat{s}_2^2}}{F_{\alpha/2;(n_1-1).(n_2-1)}};\dfrac{\dfrac{\hat{s}_1^2}{\hat{s}_2^2}}{F_{1-\alpha/2;(n_1-1).(n_2-1)}}\right)$

4. PRUEBAS PARAMÉTRICAS

Aplicación	**Tipo de contraste**	**Hip. Nula**	**Hip. Alter.**	**Estadístico K**	**Distrib.**	**Región rechazo**	**Limitaciones y observacione**
Media (1 normal)	BILATERAL (σ^2 conocida)	$\mu=\mu_0$	$\mu\neq\mu_0$	$\dfrac{\lvert\bar{x}-\mu_0\rvert}{\sigma/\sqrt{n}}$	Normal	$K>z_{\alpha/2}$	
	UNILATERAL (σ^2 conocida)	$\mu\leq\mu_0$	$\mu>\mu_0$	$\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	Normal	$K>z_{\alpha}$	
	BILATERAL (σ^2 desconocida) n>30	$\mu=\mu_0$	$\mu\neq\mu_0$	$\dfrac{\lvert\bar{x}-\mu_0\rvert}{\hat{s}/\sqrt{n}}$	Normal	$K>z_{\alpha/2}$	
	UNILATERAL (σ^2 desconocida) n>30	$\mu\leq\mu_0$	$\mu>\mu_0$	$\dfrac{\bar{x}-\mu_0}{\hat{s}/\sqrt{n}}$	Normal	$K>z_{\alpha}$	
	BILATERAL (σ^2 desconocida) n≤30	$\mu=\mu_0$	$\mu\neq\mu_0$	$\dfrac{\lvert\bar{x}-\mu_0\rvert}{\hat{s}/\sqrt{n}}$	Student	$K> t_{\alpha/2;(n-1)}$	
	UNILATERAL (σ^2 desconocida) n≤30	$\mu\leq\mu_0$	$\mu>\mu_0$	$\dfrac{\bar{x}-\mu_0}{\hat{s}/\sqrt{n}}$	Student	$K> t_{\alpha;(n-1)}$	
Varianza (1 normal)	BILATERAL	$\sigma^2=\sigma^2_0$	$\sigma^2\neq\sigma^2_0$	$\dfrac{(n-1)\hat{s}^2}{\sigma_0^2}$	Ji-cuadrado	$K\notin[\chi^2_{1-\alpha/2;(n-1)};\chi^2_{\alpha/2;(n-1)}]$	
	UNILATERAL	$\sigma^2\leq\sigma^2_0$	$\sigma^2>\sigma^2_0$	$\dfrac{(n-1)\hat{s}^2}{\sigma_0^2}$	Ji-cuadrado	$K>\chi^2_{\alpha;(n-1)}$	

Aplicación	**Tipo de contraste**	**Hip. Nula**	**Hip. Alter.**	**Estadístico K**	**Distrib.**	**Región rechazo**	**Limitaciones y observacione**
Parámetro p (1 binomial) (muestras grandes)	BILATERAL	$p=p_0$	$p\neq p_0$	$\dfrac{\lvert\hat{p}-p_0\rvert}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	Normal	$K>z_{\alpha/2}$	
	UNILATERAL	$p\leq p_0$	$p>p_0$	$\dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$	Normal	$K>z_{\alpha}$	
Parámetro λ (1 Poisson) (muestras grandes)	BILATERAL	$\lambda=\lambda_0$	$\lambda\neq\lambda_0$	$\dfrac{\lvert\overline{x}-\lambda_0\rvert}{\sqrt{\dfrac{\lambda_0}{n}}}$	Normal	$K>z_{\alpha/2}$	
	UNILATERAL	$\lambda\leq\lambda_0$	$\lambda>\lambda_0$	$\dfrac{\overline{x}-\lambda_0}{\sqrt{\dfrac{\lambda_0}{n}}}$	Normal	$K>z_{\alpha}$	

Igualdad de medias (2 normales)	BILATERAL σ^2_1 y σ^2_2 conocidas	$\mu_1=\mu_2$	$\mu_1\neq\mu_2$	$\dfrac{\lvert\overline{x}_1-\overline{x}_2\rvert}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	Normal	K>$z_{\alpha/2}$	
	UNILATERAL σ^2_1 y σ^2_2 conocidas	$\mu_1\leq\mu_2$	$\mu_1>\mu_2$	$\dfrac{\overline{x}_1-\overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	Normal	K>z_{α}	
	BILATERAL n_1 >30 y n_2 >30 σ^2_1 y σ^2_2 desconocidas	$\mu_1=\mu_2$	$\mu_1\neq\mu_2$	$\dfrac{\lvert\overline{x}_1-\overline{x}_2\rvert}{\sqrt{\dfrac{\hat{s}_1^2}{n_1}+\dfrac{\hat{s}_2^2}{n_2}}}$	Normal	K>$z_{\alpha/2}$	
	BILATERAL $n_1\leq30$,$n_2\leq30$ σ^2_1 y σ^2_2 desconocidas pero iguales	$\mu_1=\mu_2$	$\mu_1\neq\mu_2$	$\dfrac{\lvert\overline{x}_1-\overline{x}_2\rvert}{\sqrt{\dfrac{(n_1-1)\hat{s}_1^2+(n_2-1)\hat{s}_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$	Student	K> $t_{\alpha/2;(n_1+n_2-2)}$	
	BILATERAL $n_1\leq30$,$n_2\leq30$ σ^2_1 y σ^2_2 desconocidas y distintas	$\mu_1=\mu_2$	$\mu_1\neq\mu_2$	$\dfrac{\lvert\overline{x}_1-\overline{x}_2\rvert}{\sqrt{\dfrac{\hat{s}_1^2}{n_1}+\dfrac{\hat{s}_2^2}{n_2}}}$	Student	K> $t_{\alpha/2;m}$ $m=\dfrac{\left(\dfrac{\hat{s}_1^2}{n_1}+\dfrac{\hat{s}_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{\hat{s}_1^2}{n_1}\right)^2}{n_1+1}+\dfrac{\left(\dfrac{\hat{s}_2^2}{n_2}\right)^2}{n_2+1}}-2$	

Aplicación	Tipo de contraste	Hip. Nula	Hip. Alter.	Estadístico K	Distrib.	Región rechazo	Limitaciones y observacione
Igualdad de medias (2 normales)	UNILATERAL $n_1 > 30$ y $n_2 > 30$ σ^2_1 y σ^2_2 desconocidas	$\mu_1 \le \mu_2$	$\mu_1 > \mu_2$	$\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}}}$	Normal	$K > z_\alpha$	
	UNILATERAL $n_1 \le 30, n_2 \le 30$ σ^2_1 y σ^2_2 desconocidas pero iguales	$\mu_1 \le \mu_2$	$\mu_1 > \mu_2$	$\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{(n_1-1)\hat{s}_1^2 + (n_2-1)\hat{s}_2^2}{n_1+n_2-2}}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$	Student	$K > t_{\alpha;(n_1+n_2-2)}$	
	UNILATERAL $n_1 \le 30, n_2 \le 30$ σ^2_1 y σ^2_2 desconocidas y distintas	$\mu_1 \le \mu_2$	$\mu_1 > \mu_2$	$\dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}}}$	Student	$K > t_{\alpha;m}$ $m = \dfrac{\left(\dfrac{\hat{s}_1^2}{n_1} + \dfrac{\hat{s}_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{\hat{s}_1^2}{n_1}\right)^2}{n_1+1} + \dfrac{\left(\dfrac{\hat{s}_2^2}{n_2}\right)^2}{n_2+1}} - 2$	
Diferencia de Medias (2 normales) datos relacionados	BILATERAL σ_d=conocido	$d_0=0$ con $d_0=x_1-x_2$	$d_0 \neq 0$	$\dfrac{\overline{d}_0}{\dfrac{\sigma_d}{\sqrt{n}}}$	Normal	$K > z_{\alpha/2}$	
	BILATERAL σ_d=desconocido, $n>30$	$d_0=0$ con $d_0=x_1-x_2$	$d_0 \neq 0$	$\dfrac{\overline{d}_0}{\dfrac{\hat{s}_{d_0}}{\sqrt{n}}}$	Normal	$K > z_{\alpha/2}$	
	BILATERAL σ_d=desconocido, $n \le 30$	$d_0=0$ con $d_0=x_1-x_2$	$d_0 \neq 0$	$\dfrac{\overline{d}_0}{\dfrac{\hat{s}_{d_0}}{\sqrt{n}}}$	Student	$K > t_{\alpha/2;(n-1)}$	

Aplicación	Tipo de contraste	Hip. Nula	Hip. Alter.	Estadístico K	Distrib.	Región rechazo	Limitaciones y observacione
Igualdad de varianzas (2 normales)	BILATERAL	$\sigma^2_1=\sigma^2_2$	$\sigma^2_1\neq\sigma^2_2$	$\frac{\hat{s}_1^2}{\hat{s}_2^2}$	F Fisher	K$\notin [F_{1-\alpha/2;(n_1-1).(n_2-1)}; F_{\alpha/2;(n_1-1).(n_2-1)}]$	
	UNILATERAL	$\sigma^2_1\leq\sigma^2_2$	$\sigma^2_1>\sigma^2_2$	$\frac{\hat{s}_1^2}{\hat{s}_2^2}$	F Fisher	K> $F_{\alpha;(n_1-1),(n_2-1)}$	
Igualdad p (2 binomiales) (p_1 y p_2 desconocidos)	BILATERAL	p_1-p_2=D	p_1 -p_2≠D	Si D=0 $\frac{\lvert\hat{p}_1-\hat{p}_2\rvert}{\sqrt{\frac{\overline{p}(1-\overline{p})}{n_1}+\frac{\overline{p}(1-\overline{p})}{n_2}}}$ SiD≠0 $\frac{\lvert(\hat{p}_1-\hat{p}_2)-D\rvert}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	Normal	K>$z_{\alpha/2}$	$\hat{p}_1=\frac{x_1}{n_1}$ y $\hat{p}_2=\frac{x_2}{n_2}$ entonces $\overline{p}=\frac{x_1+x_2}{n_1+n_2}$
	UNILATERAL	p_1-p_2≤D	p_1-p_2>D	Si D=0 $\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\frac{\overline{p}(1-\overline{p})}{n_1}+\frac{\overline{p}(1-\overline{p})}{n_2}}}$ SiD≠0 $\frac{(\hat{p}_1-\hat{p}_2)-D}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$	Normal	K>z_α	$\hat{p}_1=\frac{x_1}{n_1}$ y $\hat{p}_2=\frac{x_2}{n_2}$ entonces $\overline{p}=\frac{x_1+x_2}{n_1+n_2}$

5. *PRUEBAS NO PARAMÉTRICAS*

Prueba de	Nombre	Escala de medida	Aplicación	Estadístico	Región de rechazo	Limitaciones y observaciones
Una muestra	Ji-cuadrado	N,O	Bondad de ajuste H_0:$F_n(x)=F(x)$	$\chi_n^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$ sigue una ji-cuadrado con g.l.=1	$\{\chi^2>cte\}$	$E_i>5$ para más del 80% N>50
	Kolmogorov-Skmirnov	O,I	Bondad de ajuste H_0:$F_n(x)=F(x)$	D=Máx.| $F_n(x)-F(x)$|	{D≥cte}	La variable en la población es continua
	Binomial	N (dicotómica)	Bondad de ajuste H_0:$F_n(x)=B(N,p)$	*- Si Np ó Nq≤ 10 p(x)= $\sum_{i=0}^{x} \binom{N}{i} p^i q^{N-i}$ *- Si Np ó Nq>10 z= $\frac{(f_m \pm 0,5) - \mu_x}{\sqrt{Npq}}$ siendo μ_x=Np	*- {T≤cte} *- {Q>cte}	Para poblaciones con 2 categorias
	Test de rachas	O, I	Aleatoriedad	*-N≤20 *- Si N>20 $z = \frac{r - E(r)}{\sqrt{V(r)}}$ donde $E(r) = \frac{2nm}{n+m} + 1 \qquad V(r) = \frac{2nm(2nm - n - m)}{(n+m)^2 (n+m+1)}$	*- $P(r \le r_1)=0,025$ $P(r \le r_2)=0,975$ .....|.......|...... *- {z≥cte}	Para datos continuos u ordinales test sobre o bajo la mediana

Prueba de	Nombre	Escala de medida	Aplicación	Estadístico	Región de rechazo	Limitaciones y observaciones
Dos muestras relacionadas	Mc.Nemar para significación de los cambios	N	Diferencia antes/despúes H_0:A=D ó A=D=0	$\chi^2 = \sum_{A,D} \frac{\left(\lvert A - D \rvert - 1\right)^2}{A+D}$ con g.l.=1 *- Si (A+D)/2<5 --> Usar dist.binomial con N=A+D, x menor de A ó D p=0.5	{χ^2>cte}	Dos categorías- Datos dicotomizados
	Prueba de los signos	O	H_0:$f_1=f_2$ H_1: $f_1>f_2$ ó $f_1<f_2$ ó $f_1 \neq f_2$	*- Si N≤ 25 $p(x) = \sum_{i=0}^{x} \binom{N}{i} 0,5^i . 0,5^{N-i}$ *- Si N>25 entonces $z = \frac{x \pm 0,5 - \frac{1}{2}N}{\frac{1}{2}\sqrt{N}}$	*- p<0.05 *- {Q>cte}	Muestras con la misma extensión
	Prueba de Wilcoxon	O,I	H_0:$f_1=f_2$ H_1: $f_1>f_2$ ó $f_1<f_2$ ó $f_1 \neq f_2$	*- Si N≤ 25; $T = \min(T_p, T_n)$ *- Si N>25 entonces $z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$	*- {T≤cte} *- {Q>cte}	

Prueba de	Nombre	Escala de medida	Aplicación	Estadístico	Región de rechazo	Limitaciones y observaciones
Dos muestras independientes	Fisher	N	H_0:Independencia $p_{ij} = p_{i.}p_{.j}$	$p_0=\dfrac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$ $p=\Sigma p_i$	{p<α} ver Tocher	Dos categorías – Datos dicotomizados. Tablas 2x2 (alguna celda $e_i<5$)
	Kolmogorov-Skmirnov	O,I,	H_0:$F_A=F_B$ H_1: $F_A>F_B$ ó $F_A<F_B$ o $F_A\neq F_B$	*- Si n=m=N≤ 40 $D_i=\left\lvert\dfrac{F_{ai}}{n}-\dfrac{F_{bi}}{m}\right\rvert$ D=máx {D_i} *- Si n=m=N>40 $\chi^2=4D^2\dfrac{n.m}{n+m}$ con g.l=2 *- también z=$\sqrt{\dfrac{m.n}{m+n}}$	*-{D≥ cte} *-{χ^2≥ cte} *-{z≥cte}	Las variables en la población son continuas
	Ji-cuadrado	N	H_0:Independencia $p_{ij} = p_{i.}p_{.j}$	d=$\displaystyle\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$ sigue una χ^2 con (r-1).(k-1) grados de libertad	*- {χ^2≥ cte}	Datos categóricos. Distinto tratamiento tablas 2x2 que tablas RxK
	Test de la mediana	O,I,	H_0: Igualdad de medianas	*- Para n_1+n_2 <20 => p=p_0 de Fisher *- Para n_1+n_2 ≥20 => $\chi^2=\dfrac{N\left(\lvert AD-BC\rvert-\dfrac{N}{2}\right)^2}{(A+B)(C+D)(B+D)(A+C)}$	*- {p≤ α} *- {χ^2≥ cte}	

Prueba de	**Nombre**	**Escala de medida**	**Aplicación**	**Estadístico**	**Región de rechazo**	**Limitaciones y observaciones**
Dos muestras independientes	U-Mann Whitney	O,I	$H_0:\mu_1=\mu_2$	*- Si n,m >20 $U_a = n.m + \frac{n(n+1)}{2} - R_a$ $U_b = n.m + \frac{m(m+1)}{2} - R_b$ se cumple $U_a+U_b=n.m$ $U=\min.(U_a;U_b)$ *- Si n,m ≥ 20 se cumple $\mu=n.m/2$ y $\sigma = \sqrt{\frac{n.m(n+m+1)}{12}}$	*- {U≤ cte} *- {Q≥cte}	
	Test de Wald-Wolfowitz	O,I,	$H_0:f_1=f_2$	*- Si n,m ≤ 20 r=rachas *- Si n,m >20 $E(R) = \frac{2n.m}{n+m} + 1$ $V(R) = \frac{2.n.m(2nm-n-m)}{(n+m)^2(n+m+1)}$	*- {r≤ cte} *- {Q≥cte}	Distribución continua. Si hay muchos pares ligados no se puede aplicar la prueba

Prueba de	**Nombre**	**Escala de medida**	**Aplicación**	**Estadístico**	**Región de rechazo**	**Limitaciones y observaciones**
K-muestras relacionadas	Q de Cochran	N	Comparación entre los K grupos. H_0: No existe diferencia entre los sujetos en los K-grupos	Q= $\frac{(k-1)\left[k\sum_{j=1}^{k} A_j^2-\left(\sum_{j=1}^{k} A_j\right)^2\right]}{k\sum_{j=1}^{k} L_j-\sum_{j=1}^{k} L_j^2}$ k=número de grupos A= suma de columnas L= suma en las filas sigue una χ^2 con K-1 grado de libertad	*- $\{\chi^2_{k-1} \geq cte\}$	
	Análisis de la varianza de Friedman	O,I	Comparación de medias H_0:$\mu_1=\mu_2=....=\mu_k$	$\chi_r^2=\frac{12}{nk(k+1)}\sum_{j=1}^{k} R_j^2-3n(k+1)$ donde n=número de filas k= número de columnas R= suma de rangos en la columna j	*- k=3 1<n<10 ó k=4 1<n<5 utilizar tablas. *-En otro caso: $\{\chi^2_{k-1} \geq cte\}$	Tamaño de los grupos iguales

Prueba de	**Nombre**	**Escala de medida**	**Aplicación**	**Estadístico**	**Región de rechazo**	**Limitaciones y observaciones**
K-muestras independientes	Ji-cuadrado	N	H_0:Independencia $p_{ij} = p_{i.}p_{.j}$	d= $\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$ sigue una χ^2 con (r-1).(k-1) grados de libertad	*- {$\chi^2 \geq$ cte}	Datos categóricos. Distinto tratamiento para tablas 2x2 que para tablas FxC
	Extensión del test de la mediana	O,I	H_0:Igualdad de medianas	d= $\sum_{i=1}^{2}\sum_{j=1}^{k}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$ sigue una χ^2 con (2-1).(k-1) grados de libertad k≡nº de columnas≡nº de grupos	*- {$\chi^2 \geq$ cte}	Formar tablas 2xK
	Análisis de la varianza de Kruskal-Wallis	O,I	H_0:No hay diferencia entre los k-grupos	- $H = \frac{12}{N(N+1)}\sum_{j=1}^{k}\frac{R_j^2}{n_j} - 3(N+1)$ N número total de sujetos y n_j el número de sujetos de cada grupo. -Si hay más del 25% de empates $H_c = \frac{H}{1-\frac{\sum T}{N^3-N}}$ donde $T=t^3-t$ y t es el número de sujetos empatados en cada grupo.	- k=3 y n_1,n_2 y $n_3 \leq 5$ tabla K-W. - En otro caso χ^2 crítico para g.l=k-1.	Las variables tienen como base una distribución continua

Prueba de	Nombre	Escala de medida	Aplicación	Estadístico	Región de rechazo	Limitaciones y observaciones
Correlación no paramétrica	C coeficiente de contingencia	N	Observar grado de causalidad	C= $\sqrt{\frac{\chi^2}{N+\chi^2}}$ donde $\sum_{i=1}^{r}\sum_{j=1}^{k}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}$ sigue una χ^2 con (r-1).(k-1) grados de libertad. r≡nº de filas≡nº de categorias k≡nº de columnas≡nº de grupos $0 \leq C < 1$	*- {$\chi^2 \geq$ cte} g.l= (r-1).(k-1)	Dato categórico. No alcanza valor 1. No es comparable. Influye frec.de la celda. Se emplea mejor: V de Cramer $V=\sqrt{\frac{\chi^2}{Nm}}$ m=min(r-1,c-1)
	Correlación de rangos de Spearman r_s	O	$r_s=0$ no hay ninguna relación	$r_s=1-\frac{6\sum_{i=1}^{N}d_i^2}{N^3-N}$ si hay ligas se utiliza la fórmula: $r_s=\frac{\sum x^2+\sum y^2-\sum d^2}{2\sqrt{\sum x^2\sum y^2}}$ donde $\sum x^2=\frac{N^3-N}{12}-\sum T_x$ $\sum y^2=\frac{N^3-N}{12}-\sum T_y$ con $T=(t^3-t)/12$ $-1 \leq r_s \leq 1$	*- Para valores N<=10 tablas de Spearman $RC_\alpha=\{r_s \geq cte\}$ *- N>10 $t=r_s\sqrt{\frac{N-2}{1-r_s^2}}$ se distribuye como t de Student de g.l=$_{N-2}$	

Prueba de	**Nombre**	**Escala de medida**	**Aplicación**	**Estadístico**	**Región de rechazo**	**Limitaciones y observaciones**
Correlación no paramétrica	Correlación de rango de Kendall:τ.	O	τ=0 no hay ninguna relación	$\tau = \frac{S}{\frac{1}{2}N(N-1)}$ si hay ligas $\tau = \frac{S}{\sqrt{\frac{1}{2}N(N-1) - T_x}\sqrt{\frac{1}{2}N(N-1) - T_y}}$ donde $T_x=1/2\Sigma t(t-1)$ y t número de observaciones ligadas en cada grupo de liga de la variable X. De forma similar T_y $-1 \leq \tau \leq 1$	*- N≤ 10 tabla de Kendall $RC_\alpha=\{S>cte\}$ *- N>10 $\mu_\tau=0$ $\sigma_\tau=\sqrt{\frac{2(2N+5)}{9N(N-1)}}$ $z=\frac{\tau - \mu_\tau}{\sigma_\tau}$ $RC_\alpha=\{z>cte\}$	
	Correlación parcial de rango de Kendall:$\tau_{xy.z}$	O	$\tau_{xy.z}=0$ no hay ninguna relación	$\tau_{xy.z} = \frac{\tau_{xy} - \tau_{zy}\tau_{xz}}{\sqrt{(1-\tau_{zy}^2)(1-\tau_{zx}^2)}}$		No tiene ninguna prueba de significación sólo observar si $\tau_{xy.z}=\tau_{xy}$ para inferir resultado por la segunda
	Coeficiente de Concordancia de Kendall: W	O	Concordancia entre k ordenaciones semejantes	*- $r_{sav} = \frac{kW-1}{k-1}$ correlación de rango de Spearman entre los $\binom{k}{2}$ pares posibles *- W=$\frac{s}{\frac{1}{12}k^2(N^3-N)}$ si hay pocas ligas. *-W=$\frac{s}{\frac{1}{12}k^2(N^3-N)-k\sum_T T}$ donde T=$\frac{\sum(t^3-t)}{12}$ siendo t=número de observaciones en un grupo ligado por un rango dado	*- n≤7 tabla de concordancias de Kendall $RC_\alpha=\{s>cte\}$ *- Si n>7 calcular $\chi^2=k(N-1)W$ -->χ^2 con gl=N-1	*- 0≤W≤1 pero r_{sav} varia entre –1 y 1

6. INTRODUCCIÓN AL MUESTREO

Muestreo aleatorio simple:

Parámetro	Estimador del parámetro	Varianza del estimador	Estimador de la varianza	Intervalo de confianza
Media	$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$	$V(\bar{y}) = \frac{S^2}{n}(1-f)$	$\hat{V}(\bar{y}) = \frac{s^2}{n}(1-f)$	$\bar{y} \pm k\sqrt{\hat{V}(\bar{y})}$
Total	$\hat{Y} = N\bar{y}$	$V(\hat{Y}) = N^2 V(\bar{y})$	$\hat{V}(\hat{Y}) = \frac{s^2}{n} N^2 (1-f)$	$\hat{Y} \pm k\sqrt{\hat{V}(\hat{Y})}$
Proporción	p= $\frac{a}{n}$ $\hat{A} = Np$	V(p)= $\frac{N-n}{N-1}\frac{PQ}{n}$ $V(\hat{A}) = N^2 V(p)$	$\hat{V}(p) = \frac{N-n}{N}\frac{pq}{n-1}$ $\hat{V}(\hat{A}) = N^2 \hat{V}(p)$	$p \pm k\sqrt{\hat{V}(p)}$ $\hat{A} \pm k\sqrt{\hat{V}(\hat{A})}$

Parámetro	Población finita	Población infinita
Media	$n = \frac{s^2}{\frac{e^2}{k^2} + \frac{s^2}{N}}$	$n = s^2 \frac{k^2}{e^2}$
Total	$n = \frac{N^2 s^2}{\frac{e^2}{k^2} + Ns^2}$	
Proporción	$n = \frac{\frac{N}{N-1} PQ}{\frac{e^2}{k^2} + \frac{1}{N-1} PQ}$	$n = \frac{k^2 PQ}{e^2}$
Caso P=Q= $\frac{1}{2}$,	$n = \frac{\frac{N}{N-1}}{\frac{4e^2}{k^2} + \frac{1}{N-1}}$	$n = \frac{k^2}{4e^2}$

Muestreo estratificado:

- Cálculo del tamaño muestral

Para la media................ n= $\dfrac{\displaystyle\sum_{j=1}^{L}\frac{W_j^2 S_j^2}{w_j}}{\dfrac{e^2}{k^2}+\dfrac{1}{N}\displaystyle\sum_{j=1}^{L} W_j S_j^2}$

Para el total................... n= $\dfrac{\displaystyle\sum_{j=1}^{L}\frac{N_j^2 S_j^2}{w_j}}{\dfrac{e^2}{k^2}+\displaystyle\sum_{j=1}^{L} N_j S_j^2}$

Para la proporción ....n= $\dfrac{\displaystyle\sum_{j=1}^{L} W_j^2 \frac{N_j}{N_j-1}\frac{P_j Q_j}{w_j}}{\dfrac{e^2}{k^2}+\dfrac{1}{N^2}\displaystyle\sum_{j=1}^{L}\frac{N_j^2}{N_j-1} P_j Q_j}$

donde $w_j = \dfrac{n_j}{n}$ y S^2 es la cuasivarianza poblacional. Recordar que $S_j^2 = \dfrac{N_j}{N_j-1} P_j Q_j$

- Afijación:

a) Afijación uniforme, donde se reparte por igual entre todos los estratos $n_j = \frac{n}{L}$

b) Afijación proporcional, la que se hace en proporción al tamaño del estrato. $\frac{n_j}{N_j} = cte.$

Teniendo, en este caso, todas las unidades muestrales la misma probabilidad de ser seleccionadas en la muestra.

c) Afijación óptima o de mínima varianza, donde se eligen los n_j de forma que minimicen la varianza para un n fijo:

$$n_j = \frac{N_j S_j}{\sum_{j=1}^{L} N_j S_j} n$$

d) Afijación para un coste, en la que se eligen los n_j de forma que minimicen la varianza para un coste fijo, C que generalmente se expresa como $c = c_0 + \sum_{j=1}^{L} n_j c_j$ donde c_j es el coste de elegir una unidad en el estrato j y c_0 es el coste inicial, se obtiene los tamaños:

$$n_j = \frac{N_j S_j \frac{1}{\sqrt{c_j}}}{\sum_{j=1}^{L} N_j S_j \frac{1}{\sqrt{c_j}}} n$$

o tambien

$$n_j = \frac{(c - c_0) W_j S_j \frac{1}{\sqrt{c_j}}}{\sum_{j=1}^{L} W_j S_j \sqrt{c_j}} n$$

Muestreo por conglomerados:

Tamaño de la muestra.- Suponiendo que se ha elegido el tamaño del conglomerado y que se intenta que sea mínima la varianza entre conglomerados, entonces el número de conglomerados para la estimación de la media poblacional será:

$$\frac{e^2}{k^2} = \frac{N-n}{Nn\bar{M}^2}.s_c^2$$

donde $s_c^2 = \frac{\sum_{i=1}^{n}(y_i - m_i\bar{y})^2}{n-1}$ que se obtiene de una muestra previa.

Y por tanto $n = \frac{Ns_c^2}{\frac{e^2}{k^2}N\bar{M}^2 + s_c^2}$